



Project no.: **045128**

Project acronym: **SAPIR**

Project title: **Search on Audio-visual content using Peer-to-peer
Information Retrieval**

Instrument: **Specific Targeted Research Project**

Thematic Priority: **FP6-2006-IST-6**

Title of report: Publishable Executive Summary

Period covered: from **1.1.2007 to 31.3.2008**

Date of preparation: **27.4.2008**

Start date of project: **1.1.2007**

Duration: **30 months**

Project coordinator name: **Yosi Mass**

Project coordinator organisation name: **HRL**

Revision **0.1**



EXECUTIVE SUMMARY

Project objectives

Web search is dominated today by search giants like Google, Yahoo or MSN that deploy a centralized approach to indexing and utilize text only indexes enriched by page rank algorithms. Consequently, while it is possible to search for audio-visual content, it is limited to associated text and metadata annotations. Supporting real content-based audio-visual search requires media specific understanding and extremely high CPU utilization which would not scale in today's centralized solutions.

SAPIR aims at breaking this technological barrier by developing a large-scale distributed P2P architecture that will make it possible to search in audio-visual content using the query by example paradigm. "A picture is worth a thousand words" so using an image taken by a cell phone to find information about e.g. a monument we bump into or singing a melody as a search hint for a full song, combined with optional metadata annotations and user and social networking context will provide the next level of search capabilities and precision of retrieved results.

Our vision is to conduct innovative research that will lead to a technology where end-users are peers that can produce audiovisual content from their mobile devices. This content will be indexed by super-peers across a scalable P2P network to enable content search in real-time while respecting IPR and protecting against spam.

To this end, SAPIR brings experts in audio-visual content understanding in the areas of text, audio, image, video and music analyzing. A common framework for feature extraction from all media contents will be developed and will be used for similarity search and ranking along all supported media.

To address scalability issues we will develop a P2P architecture where feature extraction can occur in one peer and pushed to an indexing peer. The P2P architecture will provide a scalable indexing structure that can be used for multi feature search. Caching techniques will be developed to increase system performance.

To further improve audio-visual retrieval and navigation SAPIR's consortium combines experts in Mobile devices technology with experts in Social networking and in IPR to enable a secure and trusted environment.

Having such a technology can provide significant advantage to the European community over existing centralized text only search engines and can be applied in various fields of applications such as tourism, government services, healthcare and more.

Contractors involved

- 1 IBM ISRAEL - SCIENCE AND TECHNOLOGY LTD (IBM),
- 2 ISTITUTO DI SCIENZA E TECNOLOGIE DELL'INFORMAZIONE, CONSIGLIO NAZIONALE DELLE RICERCHE (CNR)
- 3 MAX-PLANCK-GESELLSCHAFT ZUR FOERDERUNG DER WISSENSCHAFTEN E.V. (MPG)
- 4 UNIVERSITA' DEGLI STUDI DI PADOVA (UPD)
- 5 EURIX S.R.L. (EURIX)
- 6 XEROX - SAS (XRCE)
- 7 MASARYKOVA UNIVERZITA (MU BRNO)
- 8 TELEFONICA INVESTIGACION Y DESARROLLO SA UNIPERSONAL (TID)
- 9 TELENOR ASA (TELENOR)



Coordinator Contact Details

Yosi Mass (yosimass@il.ibm.com), IBM Haifa Research Lab, Haifa University Campus, Mount Carmel, Haifa 31905, ISRAEL, Tel: 972-3-768-9427, Fax: 972-3-768-9545

Work performed

This report covers the first 15 months of the SAPIR project. During this period we met all planned deliverables and milestones. We submitted 16 deliverables and we have a running demo that can be used to search in a large collection of 10M images using queries combining image and text. SAPIR partners participated in all EC concertation meetings and contributed to the roadmap of the Chorus coordination activity in the areas of P2P AV search and use cases definitions. We published about 30 related publications some of them co-authored by several partners, organized several workshops and we are actively involved and contribute to standard bodies such as MPEG and DMP.

Results achieved so far

SAPIR main objective is research and development of large scale solutions for indexing and search audio-visual content. Other objectives are support for multiple devices embedding social networking in a trusted environment. To guide all the SAPIR objectives we started by defining specific scenarios and use cases that require such large scale solutions. The outcome is deliverable D1.1 with five scenarios that will be then used for the final testbed and demonstration of SAPIR at the end of the project. Furthermore SAPIR aims at a new search paradigm combining "Query-by-example" combined with text from a range of devices so the above scenarios were tested in Focus groups and the result is deliverable D1.2 with guidelines for usability and user interface design recommendation for the above scenarios.

Search in audio-visual content is a multi-disciplinary effort that combines content analysis for extracting low level features from the audio-visual content, efficient distributed indexing structures to support similarity search, query routing, digital rights management and more. To glue all this components together a common language between people and then between software components is needed. This is manifested in SAPIR as a glossary defining the SAPIR concepts and a set of APIs. The APIs were designed to enable an architecture that is scalable (using P2P overlays), extensible (can support new media and new indexing structures) and versatile (can be used both for data centers and for autonomous peers). All this is defined in deliverable D2.1.

Our next result is a media analysis framework for extracting low level features from the different audio-visual media (image, speech, music, video and text) and representing them in a common format. We decided to use MPEG-7 and the common format is described in deliverable D3.1. The framework is based on the open-source UIMA (Unstructured Management Architecture) which allows a reuse of components such as image and speech media analyzers to analyze more complex objects such as video.

As a proof of concept we implemented the SAPIR APIs resulting in three overlays - first overlay is a P2P implementation for metric spaces and is implemented by MUFIN (Multi-feature Indexing Network) developed by MU-Brno. Second overlay is a P2P for text and is implemented by Minerva from MPI. In addition IBM implemented another text overlay based on Lucene. Those are summarized in deliverable D4.1 and D4.2.



To demonstrate the scalability of the solution we needed to have some real large data set. For this end CNR has crawled the Flickr collection and extracted over 50 million images. The images were analyzed by the SAPIR image annotator and use the SAPIR common representation schema to represent each image by five visual descriptors together with their metadata as extracted from Flickr. The result is the CoPhir (Content-based Photo Image Retrieval Test-Collection) that will be available to the scientific community for benchmarks on large scale search in audio-visual content.

A query language was defined (deliverable D5.1) and a web interface demo was built to search by using image, text or a combination of both. We showed scalability by achieving 1sec response time for a query over a collection of 1M images using 50 peers and over a collection of 10M images using 500 peers. The results show that same response time is achieved by linearly increasing the number of peer as the volume of data grows.

Work has started on protecting governed content in P2P (WP6) and on adding social networking and support for mobile devices (WP7). For protecting governed content, a study was made (deliverable D6.1) and Eurix is involved in the implementation of a reference software codenamed Chillout, of the Digital Media Project (DMP) and In particular in its extension for a P2P network. For social networking we did a study of existing technologies (reported in deliverable D7.1) and some initial experiments are conducted on implementation of self-organizing networks based on social network relations.

SAPIR partners participated in Chorus concertation meetings and in EU organized events and we contributed to the Chorus roadmap for P2P search and use cases. Eurix is responsible for the SAPIR standardization activities (Task 9.4) and is actively involved in the working groups of several standard bodies such as MPEG and DMP. SAPIR is one of the projects of FP6 invited, together with the FP7 projects in the area, to participate to the EU initiative on the "Future of Internet, perspectives emerging from R&D in Europe". SAPIR members attended the first meeting on March 31st in Bled, Slovenia, and is cooperating in drafting the white paper on "The Future Internet: A Content Creation and Media Delivery Perspective"

Expected end results

Our next steps are to show the scalability of the system by increasing the collection by another order of magnitude (50M images), using self organizing networks and demoing other media types such as speech, music and video.

To support the other objectives of SAPIR we work on interface from mobile devices, using J2ME on the device. Other on going activities are improving the search results by adding social networking information and a development of a P2P DRM solution for protecting governed content.

We have dedicated WP8 to test and integrate an end-to-end implementation of some of the scenarios defined in WP1. The prototype we develop and all our demos are based on the open P2P architecture and APIs we have defined in SAPIR thus they can be extended by 3rd parties. Moreover the SAPIR deliverables are designed to have detailed design and guidelines for developing large scale search solution for search in audio-visual content using P2P architecture.



At the end of the project a final education and exploitation plan will be published.

Plan for using and dissemination

SAPIR aims at answering a European-wide requirement: research towards the creation of new Multimedia experiences. Therefore SAPIR will address the creation and sharing of MM digital contents, enhancing both freedom and control of the digital contents and taking into account all the players in the value chain.

Exploitation of results will be done by the industrial partners and will be defined using the scenarios developed as part of WP1. The exploitation will use the SAPIR developed APIs and protocols which are reported in the SAPIR deliverables and appear on the SAPIR site. The APIs are open and extensible so can be used by 3rd parties outside SAPIR.

Exploited results so far include the CoPhIR (COntent-based Photo Image Retrieval) Test-Collection that has been developed to make significant tests on the scalability of the SAPIR project infrastructure for similarity search. CoPhIR is going to be made available to the research community to try and compare different indexing technologies for similarity search, with scalability being the key issue.

When CoPhIR will be ready for distribution, the organizations (universities, research labs, etc.) interested in building experimentations on it will have to sign the enclosed CoPhIR Access Agreement and the CoPhIR Access Registration Form, sending the original signed document to us by mail. Instructions will be then given to access and download the required files

Dissemination of SAPIR results are done through publications, participation in conferences, organizing workshops and through the Chorus and EC organized events. A Dissemination and Use plan appears on the SAPIR site at <http://www.sapir.eu/publications>.